DOCUMENT RESUME

ED 339 714 TM 017 534

AUTHOR Baker, Eva L.; And Others
TITLE The ACOT Report Card: Effects on Complex Performance and Attitude.
INSTITUTION California Univ., Los Angeles. Center for the Study of Evaluation.
SPONS AGENCY Apple Computer, Inc., Cupertino, CA.
REPORT NO CSE-TR-300
PUB DATE 89
NOTE 24p.; The appended graphs contain small type.
AVAILABLE FROM CSE Dissemination Office, UCLA Graduate School of Education, 405 Hilgard Avenue, Los Angeles, CA 90024-1521.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Academic Achievement; Achievement Tests; *Computer Uses in Education; Educational Technology; Elementary School Students; Elementary Secondary Education; Evaluation Methods; Futures (of Society); High School Students; *Outcomes of Education; Program Evaluation; *Program Improvement; Standardized Tests; *Student Attitudes; Writing (Composition)
IDENTIFIERS *Apple Classrooms of Tomorrow; School Technology Assessment Research Model

ABSTRACT

The Apple Classroom of Tomorrow (ACOT) program was designed to study what happens when "tomorrow's" educational resources, such as individual computer support, are available in the classroom. Student achievement and attitudes were evaluated at five elementary school and secondary school sites located in four states and encompassing urban, suburban, and rural communities. Data were collected during the third year of the program, using the School Technology Assessment/Research (STAR) model developed at the Center for the Study of Evaluation at the University of California (Los Angeles). The study focused on student outcomes and the appropriateness of the evaluation model. The Iowa Tests of Basic Skills, the Iowa Tests of Educational Development, the School Attitude Measure, an instrument assessing motivation and attribution, and student essays on the topic of computer use were used. The outcomes studied were standardized test achievement, performance in written composition, and student attitudes. Numbers of students varied for site and measure; data is displayed on tables. The baseline data collected provide no clear idea of ACOT success or failure, although data on achievement, writing, and attitude suggest that ACOT participation is not depriving students in any way. The findings and less formal data collection suggest potential modifications to the model, including the following: (1) reducing emphasis on local sites as the audience; (2) reviewing the feasibility of integrated cross-site data; and (3) exploring the extent to which the contexts in which ACOT is embedded create obstacles to its success. Some suggestions are made for the direction in which ACOT may evolve. A 9-item list of references and an appendix containing 45 graphs are included. (SLD)

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

J. C. BEER

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ED339714

THE ACOT REPORT CARD: EFFECTS ON COMPLEX PERFORMANCE AND ATTITUDE

CSE Technical Report 300

Eva L. Baker
Joan L. Herman
Maryl Gearhart
Center for Technology Assessment
UCLA Center for the Study of Evaluation

TM017534

BEST COPY AVAILABLE

The research reported herein was conducted with partial support from Apple Computer Company. However, the opinions expressed do not necessarily reflect the position or policy of this company and no official endorsement by this company should be inferred.

Please address inquiries to: CSE Dissemination Office, UCLA Graduate School of Education, 405 Hilgard Avenue, Los Angeles, California, 90024-1521

Introduction[1]

The Center for Technology Assessment (CT/A), a part of the UCLA Center for the Study of Evaluation, has been studying a technology-based educational program supported by the Advanced Development Group (ADG) of the Apple Computer Company. This program, the Apple Classroom of Tomorrow (ACOT), is in its third year of operation at five geographically dispersed school sites. This report documents a study that focused on student achievement and attitudes at these five sites. The research was conducted to provide baseline information for a longer term and expanded evaluation of the effects of ACOT. This report will discuss the status of ACOT, the role of UCLA-CT/A, the evaluation model, procedures, and findings.

The Apple Classroom of Tomorrow: Program Description

ACOT was originally conceived as a program to study what happens when "tomorrow's" resources, such as individual computer support, are routinely available in classrooms. Expectations for this project change, but there is a general hope among Apple personnel that computer technology will help to empower teachers and students to accomplish significant personal goals by creating a process of classroom instruction that is more interactive, cooperative, and adaptive to individual needs and aspirations. More specifically, Apple staff are interested in understanding what happens in these environments and how productive changes can occur. ACOT began in one or two classrooms at each of five sites and included classrooms of primary, upper elementary, and high school students. ADG provided computers (one at school and one at home) for each student and teacher participant. Also provided were part-time support for an on-site coordinator, access to free software, and intellectual and technical assistance, including on-site visits and conferences at Apple Computer headquarters.

Participating schools and their parent school districts sign annual contracts explicating their local responsibilities. These range from articulating local site goals, to providing appropriate electrical wiring and security arrangements for computers, to obligating teachers to prepare weekly journal entries (by voice and computer) which document the local classroom challenges. Schools also provide training for students' household members before computers are allowed to be sent home. In addition, schools also agree to participate in research and evaluation programs. On a very practical level, annual contracts may include site requests for additional equipment such as modems and printers.

The underlying philosophy of ACOT at its outset was to let each site identify how computers would be integrated substantially in its educational program. This orientation was based on the view that school innovation and improvement seem to work best when they respond to local needs and resources, in contrast to effects found for externally mandated programs. Failure of these latter programs has been attributed to the claim that they are never fully understood or accepted by the teachers and administrators who are expected to put them into action.

The ACOT program has continued to evolve in a number of ways. One major Apple decision was to permit, at some sites, the addition of classrooms to permit students' continued ACOT participation as they were promoted to higher grades, until they left the school itself. For example, a student beginning in ACOT in the 9th grade can be expected to have an ACOT environment until high school graduation. Thus, expansion has been vertical (to 10th and 11th grades) rather than lateral (additional 9th grade classrooms), and it focuses on lengthening the experience for a set of students rather than broadening the ACOT program to include different students and teachers at

---

[1] Special thanks to John Novak, Rebecca Frazier, Darlene Galuzzo, Yvonne Teruya, and Kathleen Brennan for creativity and hard work at all points of the process.

the same grade level. Another change has been the extent to which the one-student/two-computer model is central. Although at first this was a major tenet of ACOT, the ADG group is widening its approach to include laboratory and multi-media set-ups.

It should be remembered that although ACOT is an exploratory project, it is in no way a secret. Local ACOT classrooms receive and sometimes seek publicity in their hometowns and nationally in the media and at conferences. Visitors to these classrooms are common. This abnormal level of attention is important in that it itself creates a particular context for the study of ACOT. Teachers in ACOT report high levels of activity, as if they were running to keep up. In most sites, there are teachers at the same grade level without equivalent equipment and publicity and there are large numbers of students who are not in ACOT. Consequently, high expectations evolve for the ACOT project effects, both from ACOT teachers who feel they are working extraordinarily hard, from non-ACOT teachers who may feel somewhat curious or deprived, and from administrators who have made significant concessions to ACOT requirements. These expectations clearly set a context for our work.

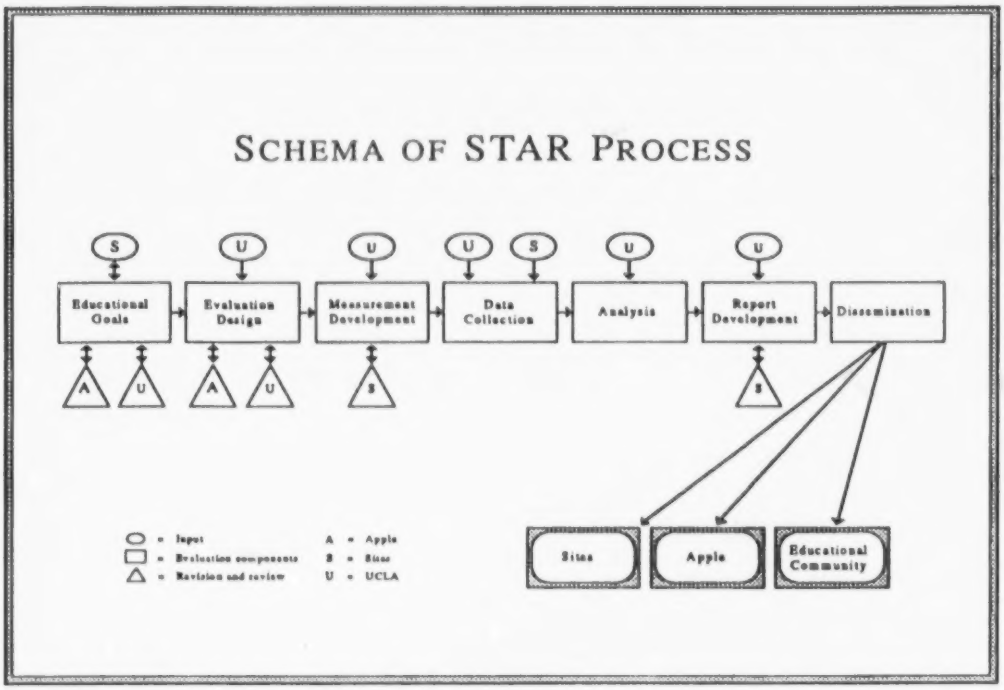
ACOT currently is being implemented in five public schools across the country. Each of these sites is unique. Together they span all grade levels from primary grades in Cupertino, California, to high school in Columbus, Ohio; their populations range from "at risk" inner-city youth in Memphis, Tennessee, to suburban children in Nashville, Tennessee, to the students of rural Blue Earth, Minnesota. Their specific goals and instructional strategies are similarly diverse.

Approach to Technology Assessment: School Technology Assessment/Research (STAR)

UCLA began formal participation in ACOT in 1987. The UCLA organizational goal was to develop a model of technology assessment appropriate for evaluating educational uses of computers and other advanced media. In addition, UCLA was interested in developing specific technological tools to assist the research and evaluation process.

The approach we are using is based on a model of technology assessment described in preliminary form elsewhere (Baker, 1987, 1988; Baker & O'Neil, 1988). The major objects of this model are the goals, processes, and outcomes of school programs. Because technology's actual use is adaptive and unpredictable, we are interested in how such goals, processes, and outcomes change over time. A second important feature of the model is the local identification and adaptation of educational goals, processes, and outcomes, all of which differ by site, grade level, and local school context. In fact, our goal of finding ways to describe and assess the ACOT program *in general*, as opposed to particular characteristics and effects of one set of classrooms in one district, is constantly at risk because the range of differences by site is so large. These differences are exacerbated by the emphasis ACOT places on local site program control and development.

Collaboration is another important characteristic of our evaluation. The schematic shown on the following page correctly represents UCLA as the initiator of evaluation processes at every point, highlights site contributions, and indicates review points. Although the contents of the evaluation report are subject to review by the sites, and we remedy corrections of fact, neither the sites nor Apple personnel have final review or editing authority.



The ACOT Evaluation Design

A detailed evaluation design was agreed upon by UCLA, Apple Computer, and the site participants (Baker & Herman, 1988). Key attributes of the plan are:

1. Collection and analysis of a broad range of potential student outcomes.
2. Collection and analysis of such information over time.
3. Linking outcome data with information on instructional process and other school context to provide explanatory power for finding.
4. Linking outcome data with demographic data to explore the extent to which ACOT differentially affects special groups of students.
5. Linking multiple indicators of key outcomes to strengthen the validity of findings.
6. Combining the strengths of both quantitative and qualitative methodologies.
7. Providing uniform data collection strategies and measures across the diverse ACOT sites, but reserving places for interests, measures, and effects unique to each site.

Identifying comparison groups for the ACOT study was a thorny issue. Randomized assignment of students was not a feature at any ACOT site, yet the desirability of some sort of comparison base, however patched up, was clear. We had

suggested identifying a comparable classroom within the ACOT school, a comparable classroom outside the school, a matched set of students from the district, or a criterion set of high-achieving students whose level of performance might set a goal for the ACOT students. Three sites attempted to comply by providing some sort of comparison sample. Two of these sites provided comparisons within the same school. Because of data quality issues, one of these comparisons is not reported. A third site provided comparison data from a neighboring school. Two sites did not comply with the request.

The 1987-88 study focused on: (a) exploring the appropriateness of measures and our design for the evaluation, and (b) using such measures to develop a profile of student outcomes as a baseline for future studies. The evaluation questions motivating the study were:

1. What are the effects on ACOT students?
2. How does ACOT affect teachers?
3. What are ACOT's effects on classroom practices?
4. What are ACOT's effects on the home?

The majority of UCLA and site time for this outcome study was devoted to the first question. Based on preliminary site visits and interviews with participating teachers, students, and administrators, the study identified a few common outcomes that were of interest across sites. Outcomes studied were student achievement on standardized tests, students' performance in written composition, and student attitudes.

Student Outcome Measures

Student Achievement

Our evaluation design called for a common norm-referenced achievement test to be administered at all sites, and the Iowa Tests of Basic Skills (ITBS) and the Iowa Tests of Educational Development (ITED) were selected for this purpose. Even though all sites routinely administered locally selected norm-referenced tests, we believed that a common measure was essential for a number of reasons. First among these was a desire for a common scale across sites to assess student progress and longitudinal growth. Although all standardized achievement tests report results in terms of standard metrics (e.g., percentile ranks, grade level equivalents, etc.), the disparities between test content and difficulty levels among the different commercial tests have been well documented. As a result, a student who scores at the 75th percentile on the California Test of Basic Skills would probably not score at exactly the same level on the Iowa Tests or on any other nationally normed test. To address this problem, and at the same time avoid unnecessary duplication of testing in the future, our plan included an approach to attempt to use the results of our common measure, the Iowa Tests, to equate locally administered standardized tests. Future data collection then might be able to use the locally administered measures rather than the Iowa Tests.

A common achievement measure also permitted additional potential benefits. In particular, we were interested in the relationship of teachers' instructional emphases and student performance. We asked teachers to indicate the level of instructional attention they provided on the particular ITBS/ITED skills assessed. Our intention was to control for reported level of instructional exposure in our data analyses.

To conserve student testing time, we selected for administration a limited number of scales from the ITBS and the ITED. Our selections, based on advice from the ITBS developers, reflected core areas of academic achievement and areas which likely

would be sensitive to ACOT effects. At the elementary school level four tests from ITBS were used: Vocabulary, Reading Comprehension, Visual Materials, and Mathematics Concepts. From the ITED tests we used the Vocabulary and Social Studies tests. The Vocabulary tests were chosen because they are generally good predictors of success in school; Reading Comprehension and Mathematics Concepts represent higher order aspects of core subject matter; the Visual Materials test was selected because of its obvious relationship to visual processing via computer display. The ITED Social Studies test focuses on higher order processes of inference and judgment in problem solving contexts.

Student Writing

All students responded to writing prompts that asked for either narrative, descriptive, or persuasive writing. These prompts were derived from those used in the International Association for the Study of Educational Achievement (IEA) Study of Written Composition (Baker, 1987). Student papers were rated by specially trained raters who used rating scales also employed in the IEA study. These scales include ratings on Overall Impression, Organization, Content, and Style.

Student Attitudes

Three approaches were used to assess student attitudes: 1) responses to a nationally normed measure (the School Attitude Measure); 2) responses to an experimental instrument that principally assesses motivation and attribution (Student Questionnaire); and 3) content analyses of students' essays on the topic of computer use.

School Attitude Measure (SAM). The SAM is a self-report survey instrument published by Scott-Foresman that consists of five scales: Motivation for Schooling; Academic Self-Concept, Performance-Based; Academic Self-Concept, Reference-Based; Sense of Control; and Instructional Mastery. Versions of these scales for Grades 4-6 and 9-12 were used in this study.

Student Motivation Questionnaire (SMQ). This instrument, developed by Stipek (1988), is based on the premise that students who believe they are competent and are provided with tasks they believe to be interesting and important will work harder in school. The instrument comprises five subscales: Perceived Competence, Intrinsic Motivation, Academic Commitment/Engagement, Belief that Ability Improves with Effort, and Value on Effort.

Essays on computer use. One persuasive essay prompt asked students to respond either positively or negatively to the idea that every student should have a computer in his school. Raters classified reasons given as either positive or negative and provided a frequency distribution of the issues most commonly raised.

Results

In our results, sites are identified by number (1-5) rather than by school or district name. Because grade level and site are confounded, the letters A and B are used to signify lower (A) and higher (B) grades at a site. For instance, at one site ninth grade would be labelled A and tenth grade labelled B, while for another site second grade would be labelled A and third grade would be labelled B. Although they make the results somewhat more difficult to read, these decisions are consequences of our agreements to protect the privacy of the teachers and students at our sites. Please note that all results reported below have excluded special education students.

Academic Achievement Results

Iowa Tests of Basic Skills/Iowa Tests of Educational Development. Figures 1.1 to 1.4 (see Appendix) present the results of the ITBS tests for three of the elementary sites. (Site 2 did not administer the tests.) The figures display mean percentile scores,[2] showing ACOT students' performance compared to the national norming sample for these tests. Because percentile ranks were generated for each grade level, the reported scores control for age and grade level differences across sites. Although two of the three sites were performing consistently above the national norm, the data verify the great differences among sites in overall student academic achievement, as well as in areas of relative strength.

Figures 2.1 to 2.4 summarize data for ACOT and comparison groups for Sites 3 and 4. Because of the baseline nature of our study, and despite the fact that students in the B groups had experienced almost two years of ACOT, we cannot assert any ACOT effects. Pre-existing differences between ACOT and non-ACOT students are unknown as is initial comparability between ACOT classes within a site.

Longitudinal achievement data from Site 3 demonstrate the nature of this interpretation problem. Figure 3 tracks the performance of one grade of ACOT students and a comparison group for three years, two years prior to and the year of ACOT exposure. The results clearly indicate initial achievement differences between the two groups that favor the ACOT cohort. Looking only at the results from the ACOT year grossly distorts effects that may be associated with ACOT. In fact, growth trends in language and mathematics did favor ACOT, but differences were more modest.

Another example of existing differences between ACOT and comparison groups can be seen in the results from Site 4. Having identified those students who would enter ACOT and those who would not for the next year, Site 4 chose to administer the ITBS to this incoming cohort during the 1988 spring data collection, providing early baseline data at this site and enabling us to investigate potential initial differences between groups. Figure 4 shows the cross-sectional results at this site for grades pre-A, A, and B. Here we see initial differences favoring the comparison group.

Figure 5 shows historical data at Site 1 for the ACOT group. Here we see that performance levels remained essentially the same. Note too that we plan to collect additional pre-ACOT performance data on ACOT and comparison samples at all sites to clarify future inferences about ACOT effects versus preexisting group differences.

Figure 6 shows results of the Iowa Tests of Educational Development for Site 5, again showing performance compared to a national norm group. No comparison group was provided for that site.

Student Writing Results

Students' essays were scored by trained raters at UCLA. All papers in the same genre (narrative, description, persuasion) were scored together, regardless of grade level. Student papers were scored in random order, and raters did not know explicitly at which grade level students were or whether students were in ACOT or comparison classes. Because most ACOT students used word processors to respond, all comparison group essays were typed prior to scoring.

All student essays were scored independently by two different raters. (A total of three raters completed the rating.) Reliability of the scoring process was assessed by computing the agreement among raters. The two different reliability indices computed for each type of essay showed very high agreement.

---

[2] The mean percentile scores for the ITBS and ITED are based on mean raw scores for each group.

Figures 7.1 to 7.4 show results from all five ACOT sites by writing type (narrative, description, and persuasion, with the computer essay analyzed separately). For purposes of brevity, only overall impression scores are provided in the body of the text. The relatively small numbers of essays from each site reflect the sampling patterns in each classroom (i.e., essay topics were randomly distributed within each classroom so that about one-third of the students wrote on each of three topics). (All students subsequently were asked to respond to the computer topic.)

Three of the sites, Sites 3, 4, and 5, reported data from two grade levels each, with letter B representing the higher grade level. Measurement was sufficiently sensitive to discriminate between lower and higher grade levels in almost every case. Interpreting these scores is problematic, in part because of the same concerns expressed about the achievement test scores reported above and in part because elementary and secondary student papers were scored together. However, the scoring scheme was criterion-referenced, meaning that each score point was referenced to a specific performance standard. A score of "3" meant competent performance. Using this interpretation, the ACOT writers appear to need additional help.

Two additional comparisons further clarify ACOT writing performance. Figures 8.1 to 8.4 compare ACOT students' performance at Grades 6 and 10 with the average score achieved by a national random sample of students at those same grade levels. ACOT students fell below the US national sample on narrative and descriptive writing. The picture was brighter in the persuasive domain: Tenth grade students were performing comparably to the national sample, while sixth grade students in Site 4 exceeded the IEA average on the persuasion/TV topic and were close on the persuasion/computer topic.

A second comparison (Figures 9.1 to 9.4) involves the ACOT groups and their local comparison classrooms for two grade levels at each of two sites. Considering grade level A only (the year of ACOT entry for any given student), the ACOT and comparison students were very closely matched in performance, with essentially equal performance for Site 3 on all four writing tasks, and a small advantage for ACOT students at Site 4. However, when one considers the B grades, where ACOT students experienced two years in the program, the direction of the differences favored ACOT students in seven out of eight cases, with differences between ACOT and comparison sites ranging from a half to a full standard deviation. Such results could have occurred because of an inaccurately matched comparison group. More positively, they might have resulted from extended ACOT exposure and the additional writing experience that exposure may facilitate. Clearly, studying the trends in differences between ACOT and comparison students in writing next year will allow us to determine whether ACOT students increase their advantage.

Among explanations for the overall level of performance is a potential mismatch between writing instruction and the structure of the writing assessment. A number of sites reported strong emphasis on a writing process consisting of planning, writing, revision, and editing over several instructional periods. The writing assessment, in contrast, asked students to compose and produce essays in a single period. Although the IEA students were similarly disadvantaged by our writing measure (which required students to compose an essay on the spot), a measurement strategy that mirrored a process approach to writing might be warranted, particularly if teacher "help" at the time of measurement could be controlled.

Attitude Measurement Results

Student responses to the SAM are reported in terms of national percentile scores, where 50 represents the score point below which half of the national norm

sample fell.[3] (See Figures 10.1 to 10.5.) Notice that the ACOT students' attitudes were predominantly more positive than the national average for all sites except Site 2.

When contrasted with their local comparison sites (Figures 11.1 to 11.5), ACOT results appear promising. ACOT students look considerably more positive for A and B grades at Site 4, and for grade A at Site 3.

Table 1 presents the results from the newly developed Student Motivation Questionnaire. Students responded to this instrument on a five-point scale, where for each item a response of 5 was equivalent to very positive, 3 equivalent to neutral, and 1 equivalent to very negative. Items making up each scale were then totaled and averaged to arrive at subscale scores. Often because of the nature of such self-report instruments, a response average of approximately 3.5 or above is considered positive and those below 2.5 are considered negative. However, because this is a new experimental measure, such rules of thumb must be used cautiously.

Results show considerable diversity in patterns across sites, with the exception that students at all ACOT sites were relatively most positive in their belief that ability can improve with effort. Interestingly, all of the ACOT groups save one grade at one site scored positively on three of the five subscales. ACOT students also were generally positive about their feelings of competency.

Comparisons between ACOT and comparison students (see Table 2) showed similar patterns in attitudes to those reported for the Student Attitude Measure (SAM). Students from Site 3, grade A, and both grades at Site 5 generally showed more positive attitudes than their comparison groups.

The content analysis of students' essays on whether all students should have computers showed overwhelming support for the affirmative. Students on average described about five benefits associated with computers for every negative aspect mentioned. Among the most common benefits students cited were: the educational value of computers (computers help me learn more and learn faster); motivational and entertainment value (computers make school and learning fun, let me learn while playing); and the career development value (computer knowledge will help me advance, will open job opportunities). Almost two-thirds of the positive comments fell into these categories. Of the reasons given for not giving computers to all students, concerns about cost and fear of vandalism accounted for 70% of students' responses. Patterns of responses were similar for both ACOT and comparison students and surprisingly similar across grade levels.

Implications of the Results

In this section, we highlight the implications of our work for understanding the effectiveness of ACOT, for planning continuing data collection, for adapting our model of technology assessment, and for improving the concept and operation of the ACOT program itself.

ACOT Effectiveness

The overall effectiveness (or conversely the unworthiness) of the ACOT program cannot be inferred from the outcome results we have presented because of design and measurement constraints. The heart of the design issue is whether differences can be ascribed to the ACOT intervention, to preexisting differences in students, or to some other factors. For example, we do not know whether differences in grade-level performance at ACOT sites were caused by ACOT participation or normal maturation, or

---

[3] The mean percentile scores for the SAM are based on mean weighted raw scores for each group.

merely reflected preexisting differences in the cohorts of students beginning their participation in 1986-87 and 1987-88. We have similar problems interpreting comparative data from non-ACOT students for those sites that provided such data. It is true that almost all evaluation studies without randomized assignment to programs suffer from similar ambiguities, but some uncertainty is reduced when multiple classrooms are involved at a single grade level. In this study we are dealing with very small numbers of classrooms, usually only one at a particular grade level. This fact makes inferences about ACOT confounded with judgments about the teachers' influences, curriculum selection, school ambiance, and the characteristics of students, among other lesser matters. Nonetheless, comparisons to local site "control" groups or to nationally reported norms are essential, even though the information they provide resolves only some of the uncertainty.

Constraints on measurement create another limitation on our data. We urged a common achievement instrument on our sites because of the recognition that local districts use different tests and even may change tests for different grades within the district. In our quest for common or cross-site inferences, getting all sites on the same measurement base seemed to be a minimum requirement. At best, we would be able to compare the sites with students at common grade levels(e.g., fourth grade). We also hoped to be able to conduct an equating study that would enable us to convert performance of varying local district tests to a common scale. Such estimates would allow us to reduce the data burden on ACOT students and especially comparison groups. Not all sites complied with our requests, and considerable resistance to such measures has developed at more than one site.

Even with full compliance on a common achievement measure, potential problems remain in the lack of congruence between such measures and local site objectives. In an attempt to accommodate our evaluation to such criticism and to increase the validity of our interpretation of the Iowa Tests, we asked teachers to complete an instrument reporting their emphases on items on the Iowa Tests. As noted earlier, we planned to use such teacher reports to weight student achievement. Because of timing and other constraints, an insufficient number of sites provided these data, so our plan could not be implemented. The consequence was that interpretability of those results was reduced.

A second effort at cross-site comparability involved the use of the IEA writing measures. We have somewhat higher hopes for such measures since almost all ACOT sites cite the importance of their writing program. We plan to continue such measures in the future.

Our third major source of data was the attitude scales dealing with issues of motivation, responsibility, and student self-concept. Almost all measurement experts have great complaints about such measures, complaints that are focused on their usefulness and reliability. Since anecdotal commentary of most technology-based innovations emphasizes student excitement and interest, omitting such measures from our evaluation would be unthinkable. Yet, their interpretation is another matter. What are we to make of relatively high self-concept for academic learning, for example, and relatively poor achievement? How should we expect such measures to correlate and should those correlations strengthen over time?

The recognition of the limitations of outcome information led us in the direction of careful process assessment of ACOT classrooms. At the present time, we are developing new classroom observation protocols that focus on critical attributes of learning, and, in a sense, add to the usual issues of classroom organization and discourse the extent to which the classroom (and computer-based instruction) provides specific opportunities to stimulate students' deeper cognitive processes. Ideally, we should take this instrument to comparison classrooms as well.

A last point should be made. Compared to the routine inferences made about classroom and school achievement based on mandated district and state assessments, ACOT evaluation procedures hold up pretty well. It is only when we probe beneath the numbers for their meanings that we confront problematic choices in design and assessment.

Implications for our Technology Assessment Model

What are the implications of our results and interpretations for the model we have developed for technology assessment? The answer has at least three components. First, our model was based upon meeting the needs of the widest audience possible. In our original design (Baker, 1987) we posited reporting to the Apple Computer Advanced Technology Group, to the local site personnel, including teachers, coordinators and district managers, and to the educational professional and research communities at large. The nature of the report was to be principally formative, focused on the improvement of the program, for Apple and ACOT personnel, and shaded more toward summative evaluation and dissemination of technology assessment models, tools, and procedures for the larger educational communities. The more we get involved, the less secure we are that we can provide useful and sufficiently detailed information for the sites to use systematically in the improvement of their local educational efforts. Certainly this option cannot be foreclosed, but the diversity of ACOT implementation, localities, personnel, and students, and the diffuseness of their goals reduces the likelihood that our findings will have direct application into instructional programs.

The obstacles to such ready incorporation of data into new plans are many, beginning with well-known cautions about utilization (see Weiss, 1972; Alkin, 1979). More directly, however, the time and intensity of on-site involvement required by the evaluators for detailed program contribution would cost far more than the budget available. And although the present evaluation model has strong interactive characteristics, our involvement is on a periodic rather than an extended basis. Because of these and other less well formulated notions, it is likely that the utility of our findings to sites will be somewhat restricted, swamped as they are by other formal, authoritative evaluation requirements of more salience to local districts and parents. In addition, the daily involvement with ACOT creates its own beliefs, and their credibility is likely stronger than annual data reported by distant experts. Even though we now believe it appropriate to consider sites as secondary audiences for our data, we will struggle to maintain the validity, credibility, and usefulness of our information to them.

In addition to audience, a serious issue for us is the extent to which we can continue to conceive of this evaluation in a cross-site context. For one thing, the site diversity, ranging from primary to high school students, makes common core measurement difficult. Another concern is the likelihood that, even if ACOT were successfully implemented, we would expect greater diversification in goals and methods across sites to adapt to local resources, objectives, and students, with the obvious reduction in those aspects that are common across sites. This is the way that technology-push (Glennan, 1967) works: The very existence of the technology creates diverse options unforeseen by developers. This centrifugal force works in the direction of reducing common attributes of separate implementations. It should be clear that the present model is based on a relatively lean and highly general set of common measures: a common set of standardized achievement measures, writing measures, attitude measure, classroom process measure, teacher questionnaire, and parent survey. But we are aware that even this minimalist set may be too procrustean for the diversity of the sites.

A third and critical point is the extent to which the STARFRAME model, focused on the objects of schooling—goals, processes, and outcomes—misses a huge source of the action. We have only begun to understand how the school and district context in which ACOT sites are embedded seriously affects the nature of the program as well as

the validity of our data collection. Based more on anecdotal, informal conversations than on objective evidence, we have become aware that ACOT classrooms themselves create organizational dynamics that vary from place to place and may be as important as the attempts at innovative instructional programs under development. Classified in our model as unanticipated outcomes, such consequences may suggest a more systematic study of how ACOT classrooms relate to their organizational homes.

Let's consider some hypothetical conditions. Imagine that at one site the principal compensates non-ACOT classrooms with all discretionary resources, even those normally available to the pre-ACOT environment. What is the ACOT program in this school and to what should it be compared? Imagine that at another site the non-ACOT teachers feel competitive and comparatively undernourished. What if they were to isolate ACOT teachers and talk negatively to their students about the ACOT program? What might be the effects on ACOT teachers and children, and on the school? What if ACOT students stay together as they progress up the grade levels? Consider the view other students may have of them and ways that students may attempt to compensate for their "difference." What about the effect of attention on students and teachers? Could students feel pressured to perform especially well because they have had the "advantage" of ACOT? How can teachers deal with day-to-day setbacks, with the sheer press to try something new because of the high expectations set for them by the district and the site manager?

Although most of these context examples have been negative, we have imagined another way to imagine the ACOT-context interaction, and this one is positive. Work by Bryk and Driscoll (1987) that manipulates large data sets has hypothesized the importance of the school community as a factor that has long-term implications for student achievement and teacher satisfaction. Community was conceived of consisting of factors such as shared goals and activities rather than simply positive school climate. Perhaps the flip side to the implicit conflicts that a highly resourced ACOT-like program may create in a school is the sense of community developed among ACOT teachers and staff and among them and their students. Such a community can exist only if there is a critical mass of adults in the environment. The sense of community may extend beyond the particular sites themselves as well to other ACOT teachers and staff who confront similar problems.

This community development strikes us as an interesting area to pursue in the ACOT studies for two reasons. First, if such community develops (and we can measure it using the same instruments as Bryk and Driscoll), it may be a shorter-term outcome presaging higher and sustained achievement in the future. Second, community is probably worth developing—as an end in itself—even if it has no discernible effects (positive or negative) on student outcome measures. We plan to explore this issue informally this year, to evaluate its utility in our model and in our measurement base, and to incorporate its systematic measurement, if warranted, next year. Such concerns seem to warrant some level of study in order to help us understand larger causes and consequences for our findings. Our model will be adapted to incorporate all three areas discussed—audience, cross-site comparison, and context—as indicated by results in the future.

Implications of Our Study for ACOT Program Improvement

After all the hedging on the interpretability of this year's outcome data, we would be foolish to make ironclad recommendations for the modification of ACOT. Nonetheless, we have acquired extensive information through objective and informal means by visiting sites, by talking with teachers and research coordinators, with school district personnel, and with ATG and other Apple staff, and by reading ACOT site plans and the research of our colleagues. We have a list of speculative issues that, if confirmed by other data or judgment, could influence changes in direction of the ACOT program. These are presented as they are relevant *now* at this point of ACOT

development, and are not meant to second-guess the design or conception that motivated ACOT and its participants at the start.

1. None of us will dispute that enormous effort has been expended by teachers and site coordinators (and by ATG staff) on trying to make ACOT work. But because of the early notion that computers should be central to *all* major aspects of schoolwork, a frenetic and partly unfocused level of activity ensued. Even though there has been considerable attempt by ATG staff to have local sites develop focused objectives, those produced to date are fairly general and skip between plans that have a direct impact on student performance and those that seem to be closer to attempts to comply with the breadth of computer integration that the original concept of ACOT seemed to demand. Perhaps ACOT sites should cut back and focus on only one or two areas and direct their considerable energy on more limited targets.

2. Even if sites are successful in identifying focused goals and delimiting what ACOT tries to achieve, we believe that some sites may desire (or need) additional substantial technical assistance in implementing an effective program for these goals. The fact is that all teachers have differential strengths and weaknesses; their training and experience may range from excellent to poor; the curriculum they are supposed to implement may be sound or uneven; the textbooks and other materials they have are likely to be insufficient guides for pedagogy. In other words, it may be time to modify the bottom-up notion that local sites should explore the use of computers and develop some specific assistance programs to help ACOT become effective.

3. Should context emerge as an important component in our assessment model, it will do so because it is an important determiner of ACOT success. The present sites operate almost exclusively as classrooms located in non-ACOT settings; such settings appear to differ considerably in terms of their hospitality to the ACOT programs. An alternative model might involve developing an ACOT school at the elementary level, and including every teacher at all grade levels so that a spirit of cooperation (or even community) is fostered and supported by all staff and leadership. At the high school level, a school within a school (an idea spawned during the counterculture sixties) might be feasible. Although it is a noble idea to drop ACOT into typical school environments and see what happens, at some point ATG might consider either this alternate model or additional models of how tomorrow develops in schools (although the Apple School of Tomorrow acronym is hardly euphonious).

ACOT has taken on the toughest set of requirements: real schools, real districts, a wide range in community characteristics and expectations, in grade levels, teacher experience and expectations, administrative leadership, parental support, and more. It is the messy real world, with personality and politics playing their usual roles. Perhaps it is our university perspective that wishes for a less complex world or at least less complex contexts in which to study ACOT. But questions like the following are hard to avoid. How many resources should be devoted to studying how technology works in such contexts? Should there be an attempt to reduce some of the variation? Should any ACOT resources be devoted to school-level development? Or to demonstrating that a specific ACOT program, under ideal conditions, results in any of the following: in better achievement of existing objectives, acquisition of new and wonderful concepts and skills, and the development of motivated learning? The ACOT sites right now may be in the process of providing such a proof-of-concept. But ATG should consider the lessons of parallel R & D, and may wish to add eggs to at least one other basket.

Summary

The baseline data collection provides no clear notion of ACOT success or failure, although the data on achievement, writing, and attitude suggest that ACOT participation is unlikely to be depriving students in any way. The difficulty in interpretation derives from the definition of comparison groups, the few numbers of participating classes, and the range of classroom contexts in which ACOT has been implemented. We believe these limitations will be remedied at least partly by longitudinal data collection and by extended classroom observations.

The implications of these findings for next year's effort involve the development of new measures of problem solving and writing (to more closely parallel local site emphases) and a systematic study of classroom processes, of teacher instructional emphases and reported stress, and of parent involvement.

These findings and other less formal data collection suggest potential modifications in our model, including reducing the emphasis on local sites as audience, reviewing the feasibility of integrated, cross-site data, and exploring the extent to which the contexts in which ACOT is embedded create obstacles to its success. In addition, we are interested in exploring the extent to which the ACOT experience develops communal spirit among its participants, and if so, how that community influences achievement.

Finally, we have made some probably provocative and, at present, certainly unsubstantiated guesses about how ACOT might evolve, including a suggestion of restricted focus, technical assistance on program development, and an entire school made up of ACOTs. These suggestions seem natural, given where ACOT is now, and are directed toward increasing the probability that any positive impact that technology may have on children and schools will be detected.

References

Alkin, M.C. (1979). *Using evaluations: Does evaluation make a difference?* Beverly Hills, CA: Sage Publications.

Baker, E.L. (1988). *Sensitive, sensible evaluation of ACOT* (Report to Apple Computer). Los Angeles: UCLA Center for the Study of Evaluation.

Baker, E.L. (1987). *Time to write: The U.S. Study of Written Composition.* Los Angeles: UCLA Center for the Study of Evaluation.

Baker, E.L., & Herman, J.L. (1988). *Implementing STAR: Sensible technology assessment/research* (Report to Apple Computer). Los Angeles: UCLA Center for the Study of Evaluation.

Baker, E.L., & O'Neil, H.F. (1988). *The evolution of the ACOT evaluation model* (Report to Apple Computer). Los Angeles: UCLA Center for the Study of Evaluation.

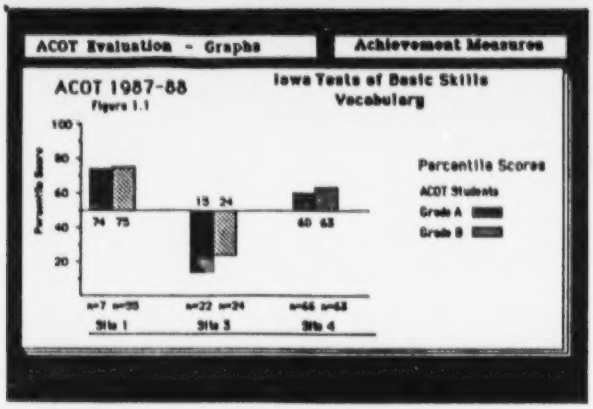
Bryk, A.S., & Driscoll, M.E. (1987). *The high school as community: Contextual influences and consequences for students and teachers.* Madison, WI: University of Wisconsin-Madison, National Center on Effective Secondary School.

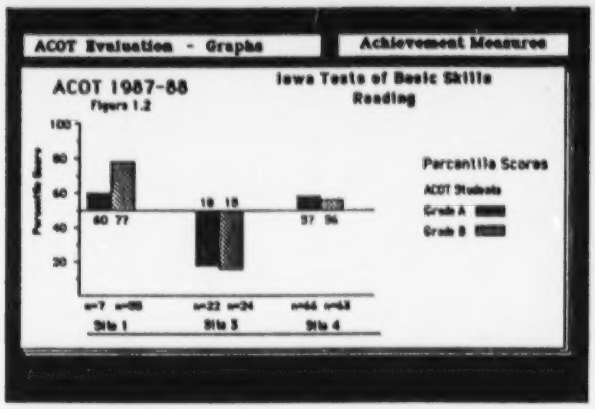
Glennan, T.K. (1967). Issues in the choice of development policies. In T. Manschak, T.K. Glennan, & R. Summers (Eds.), *Strategies for research and development.* New York: Springer-Valdez.

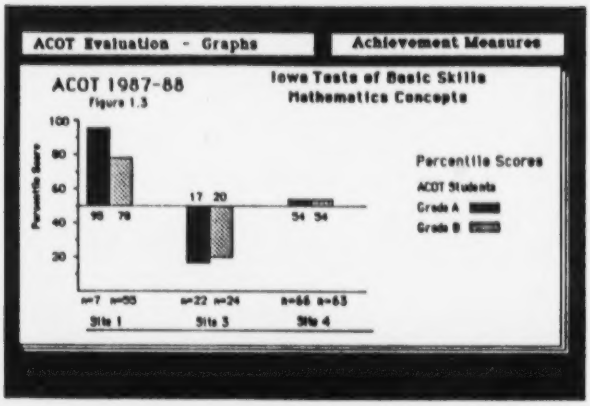
Stipek, D. (1988). *Student Measurement Questionnaire* (Unpublished testing materials). Los Angeles: UCLA Graduate School of Education.

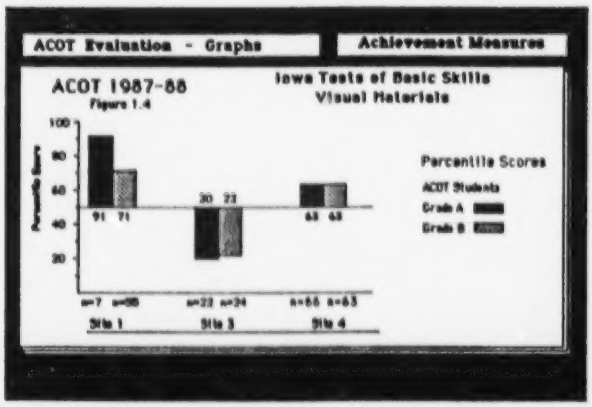
Weiss, C.H. (1972). *Evaluating action programs: Readings in social action and education.* Boston: Allyn and Bacon.

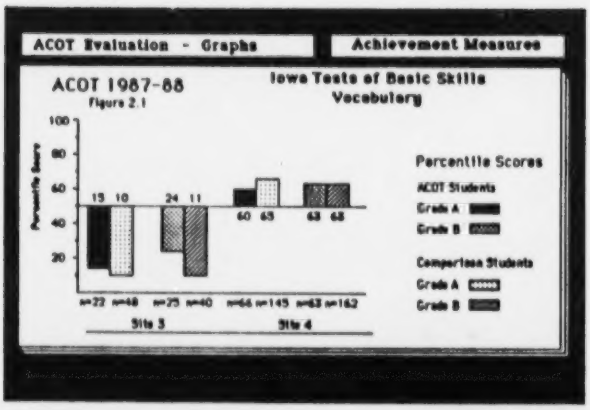
Appendix

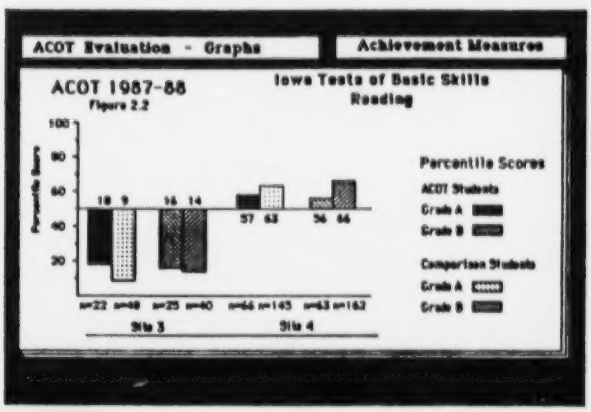




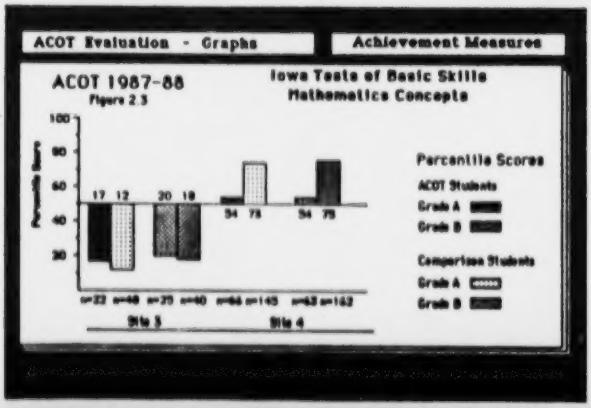


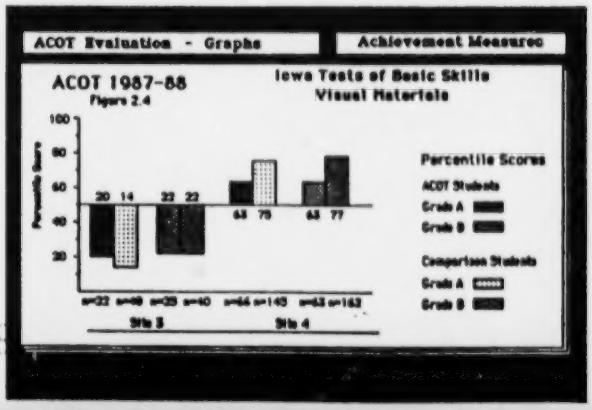


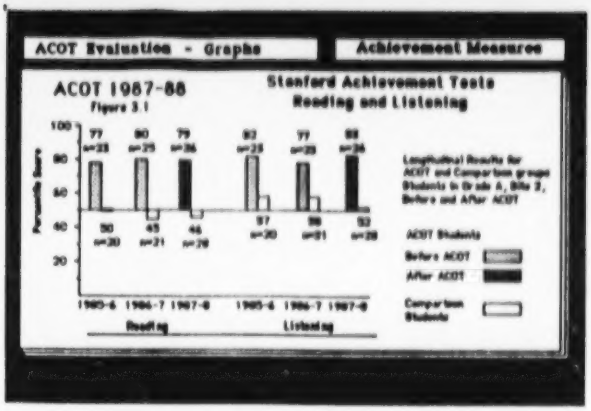


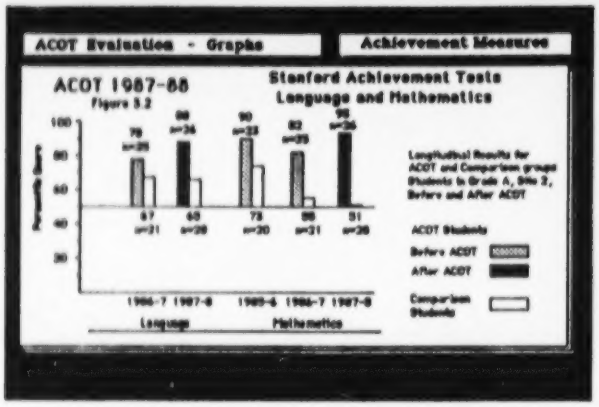


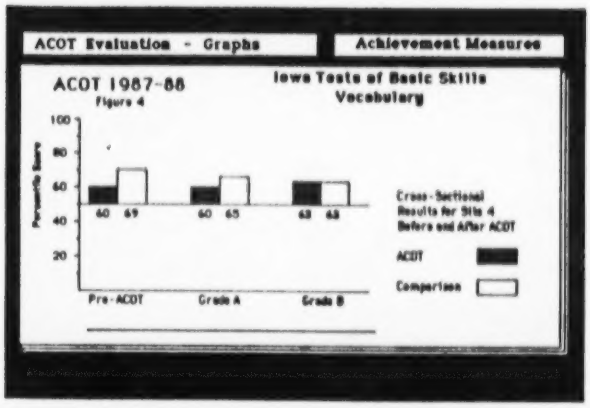
BEST COPY AVAILABLE

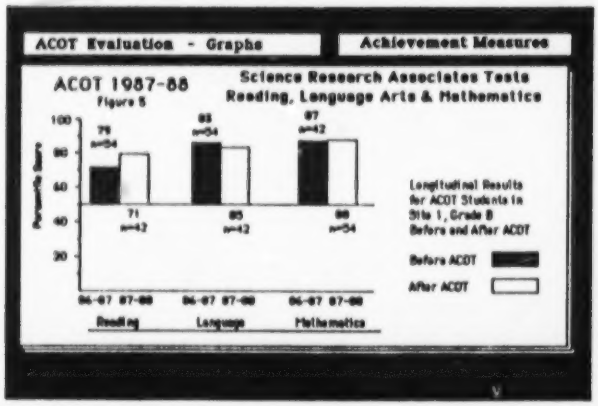


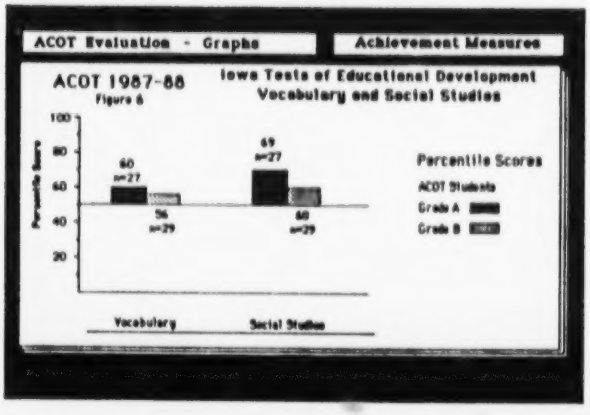


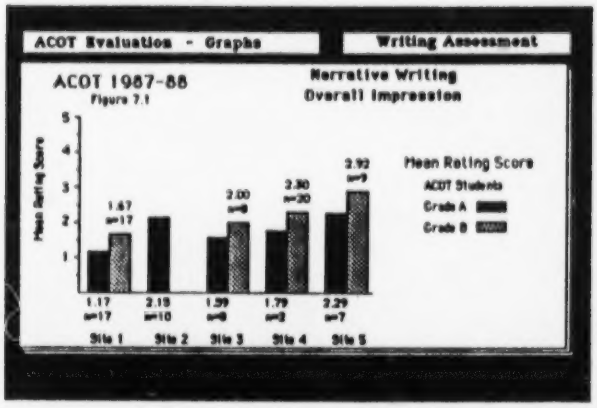

ACOT Evaluation - Graphs
Achievement Measures
ACOT 1987-88
Figure 3.1
Stanford Achievement Tests
Reading and Listening
Percentile Score
Reading
Listening
ACOT Students
Before ACOT
After ACOT
Comparison Students

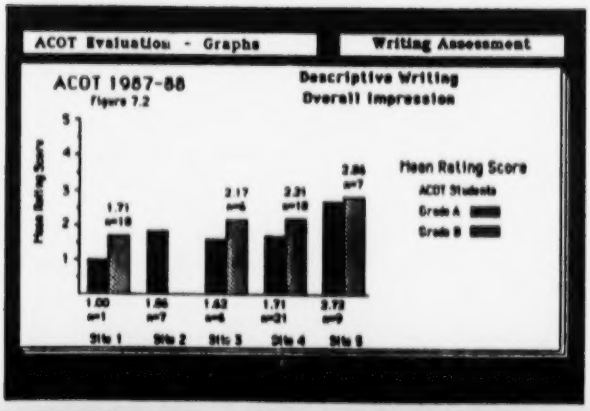

ACOT Evaluation - Graphs
Achievement Measures
ACOT 1987-88
Figure 3.2
Stanford Achievement Tests
Language and Mathematics
Percentile Score
Language
Mathematics
ACOT Students
Before ACOT
After ACOT
Comparison Students

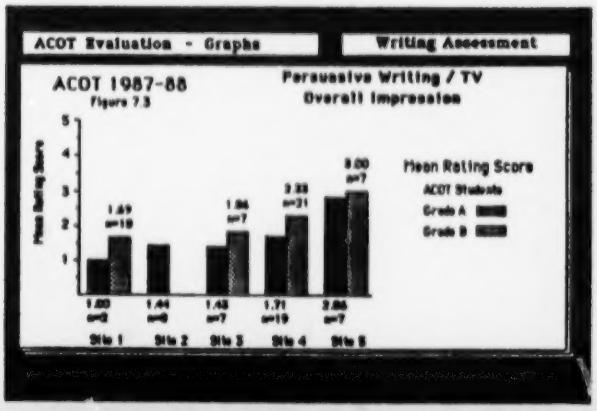

ACOT Evaluation - Graphs
Achievement Measures
ACOT 1987-88
Figure 4
Iowa Tests of Basic Skills
Vocabulary
Percentile Score
Pre-ACOT
Grade A
Grade B
Cross-Sectional Results for Site 4 Before and After ACOT
ACOT
Comparison

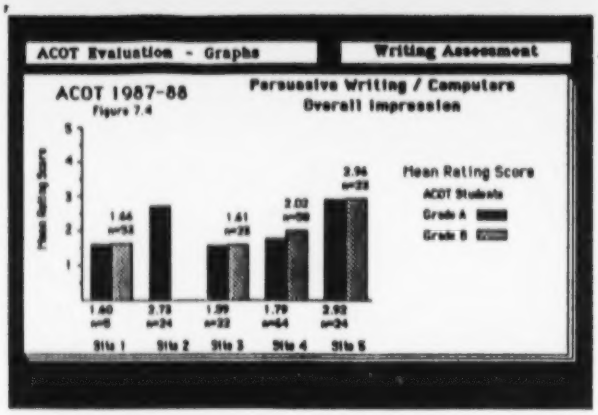

ACOT Evaluation - Graphs
Achievement Measures
ACOT 1987-88
Figure 5
Science Research Associates Tests
Reading, Language Arts & Mathematics
Percentile Score
Reading
Language
Mathematics
Longitudinal Results for ACOT Students in Site 1, Grade B Before and After ACOT
Before ACOT
After ACOT

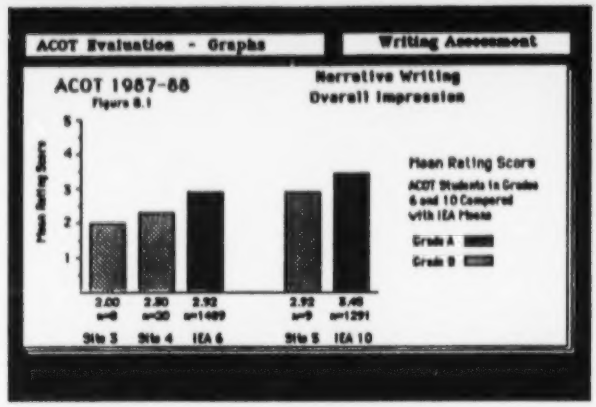

ACOT Evaluation - Graphs
Achievement Measures
ACOT 1987-88
Figure 6
Iowa Tests of Educational Development
Vocabulary and Social Studies
Percentile Score
Vocabulary
Social Studies
Percentile Scores
ACOT Students
Grade A
Grade B

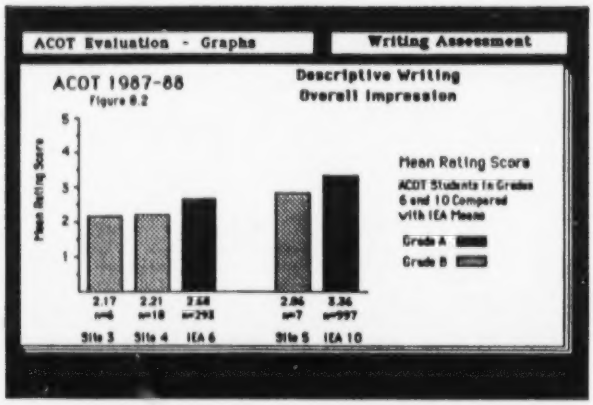

ACOT Evaluation - Graphs
Writing Assessment
ACOT 1987-88
Figure 7.1
Narrative Writing
Overall Impression
Mean Rating Score
Site 1
Site 2
Site 3
Site 4
Site 5
Mean Rating Score
ACOT Students
Grade A
Grade B

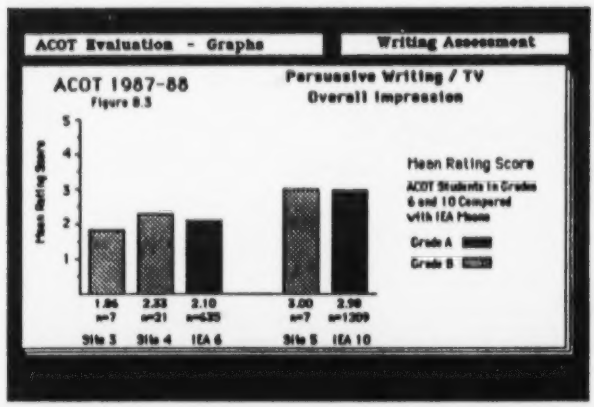

ACOT Evaluation - Graphs
Writing Assessment
ACOT 1987-88
Figure 7.2
Descriptive Writing
Overall Impression
Mean Rating Score
Site 1
Site 2
Site 3
Site 4
Site 5
Mean Rating Score
ACOT Students
Grade A
Grade B

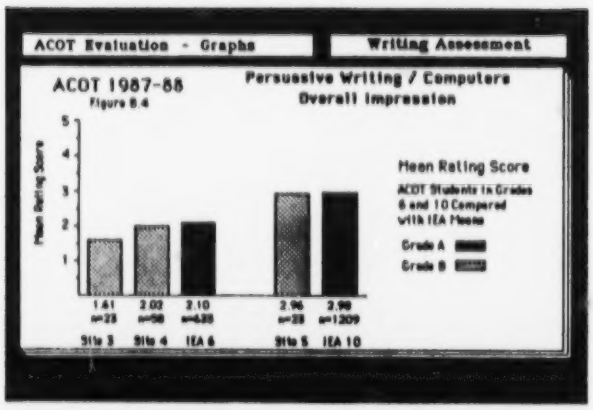

ACOT Evaluation - Graphs
Writing Assessment
ACOT 1987-88
Figure 7.3
Persuasive Writing / TV
Overall Impression
Mean Rating Score
Site 1
Site 2
Site 3
Site 4
Site 5
Mean Rating Score
ACOT Students
Grade A
Grade B

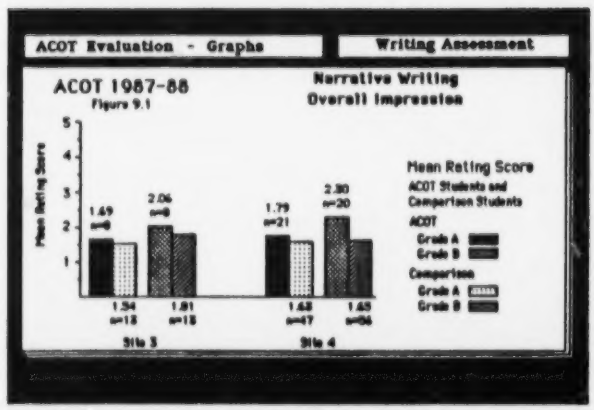

ACOT Evaluation - Graphs
Writing Assessment
ACOT 1987-88
Figure 7.4
Persuasive Writing / Computers
Overall Impression
Mean Rating Score
Site 1
Site 2
Site 3
Site 4
Site 5

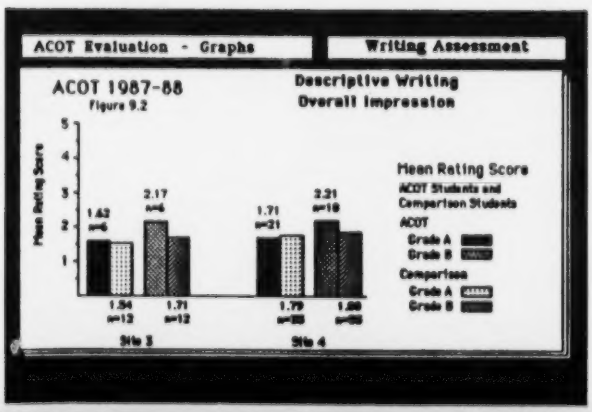

ACOT Evaluation - Graphs
Writing Assessment
ACOT 1987-88
Figure 8.1
Narrative Writing
Overall Impression
Mean Rating Score
Site 3
Site 4
IEA 6
Site 5
IEA 10

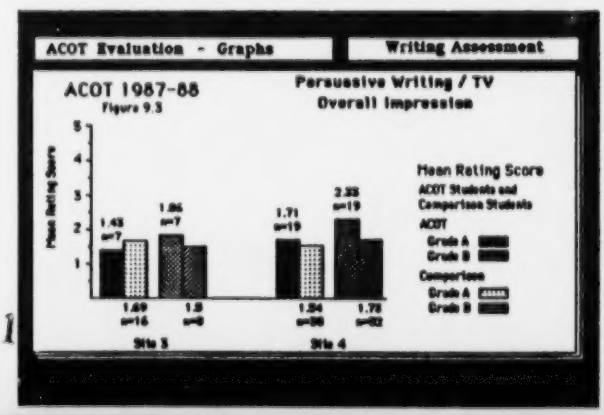

ACOT Evaluation - Graphs
Writing Assessment
ACOT 1987-88
Figure 8.2
Descriptive Writing
Overall Impression
Mean Rating Score
Site 3
Site 4
IEA 6
Site 5
IEA 10

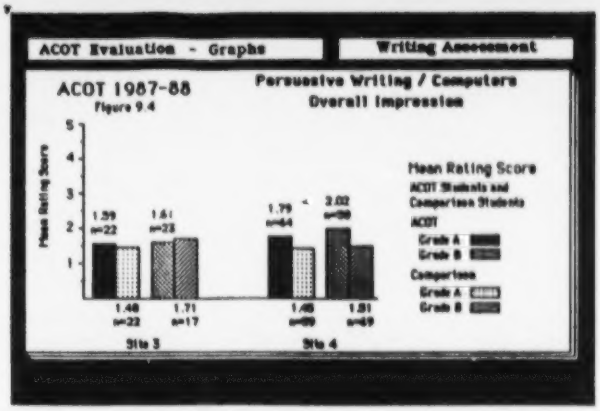

ACOT Evaluation - Graphs
Writing Assessment
ACOT 1987-88
Figure 8.3
Persuasive Writing / TV
Overall Impression
Mean Rating Score
Site 3
Site 4
IEA 6
Site 5
IEA 10

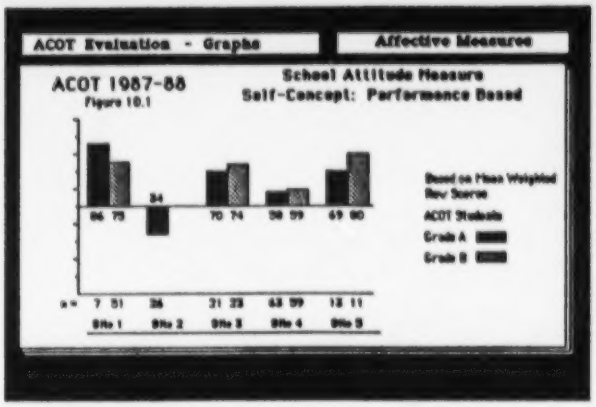

ACOT Evaluation - Graphs
Writing Assessment
ACOT 1987-88
Figure 8.4
Persuasive Writing / Computers
Overall Impression
Mean Rating Score
Site 3
Site 4
IEA 6
Site 5
IEA 10

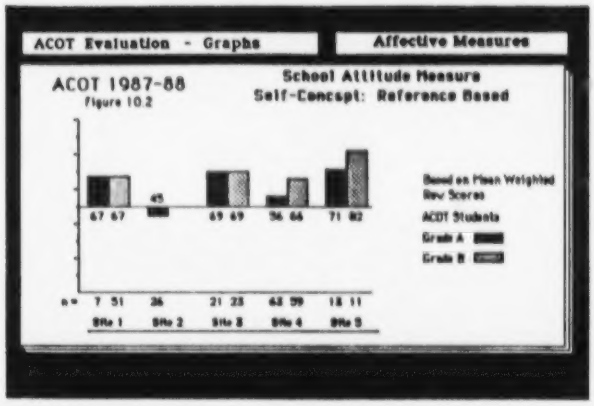

ACOT Evaluation - Graphs
Writing Assessment
ACOT 1987-88
Figure 9.1
Narrative Writing
Overall Impression
Mean Rating Score
Site 3
Site 4

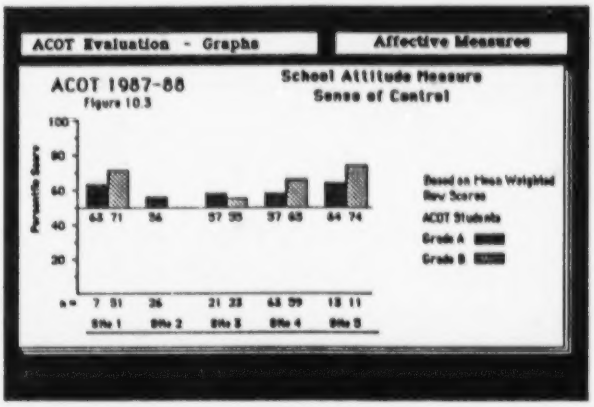

ACOT Evaluation - Graphs
Writing Assessment
ACOT 1987-88
Figure 9.2
Descriptive Writing
Overall Impression
Mean Rating Score
Site 3
Site 4

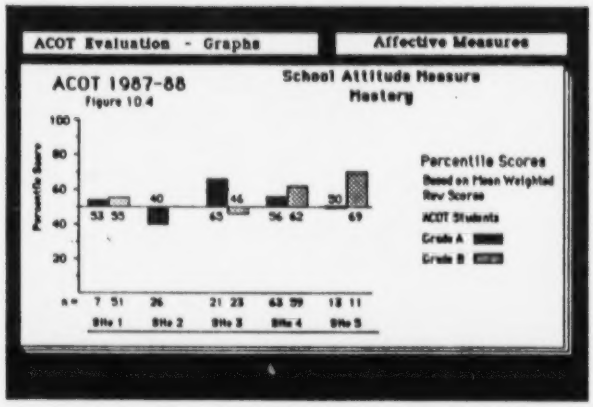

ACOT Evaluation - Graphs
Writing Assessment
ACOT 1987-88
Figure 9.3
Persuasive Writing / TV
Overall Impression
Mean Rating Score
Site 3
Site 4

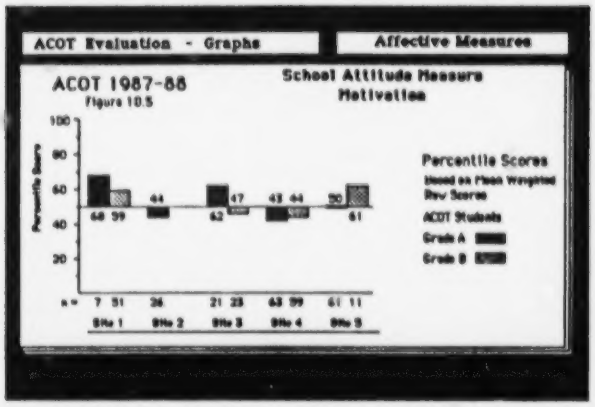

ACOT Evaluation - Graphs
Writing Assessment
ACOT 1987-88
Figure 9.4
Persuasive Writing / Computers
Overall Impression
Mean Rating Score
Site 3
Site 4

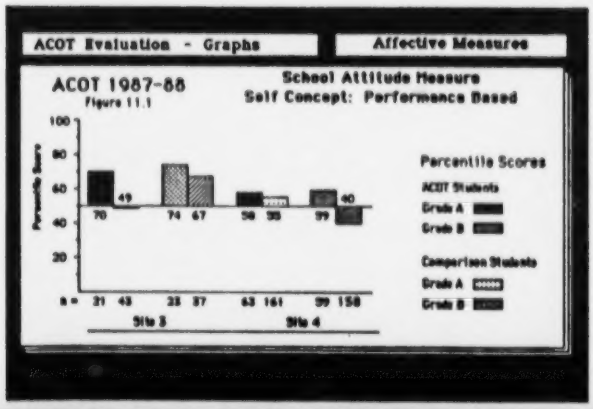

ACOT Evaluation - Graphs
Affective Measures
ACOT 1987-88
Figure 10.1
School Attitude Measure
Self-Concept: Performance Based

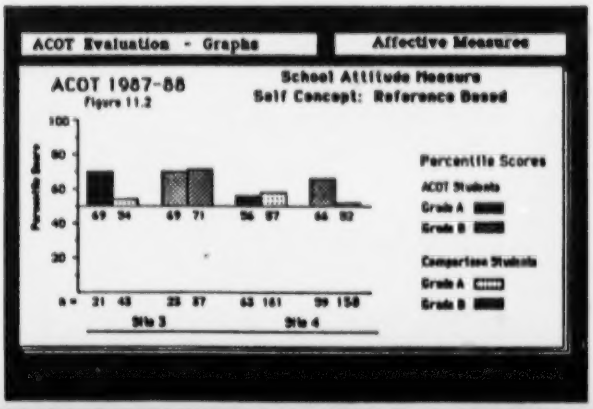

ACOT Evaluation - Graphs
Affective Measures
ACOT 1987-88
Figure 10.2
School Attitude Measure
Self-Concept: Reference Based

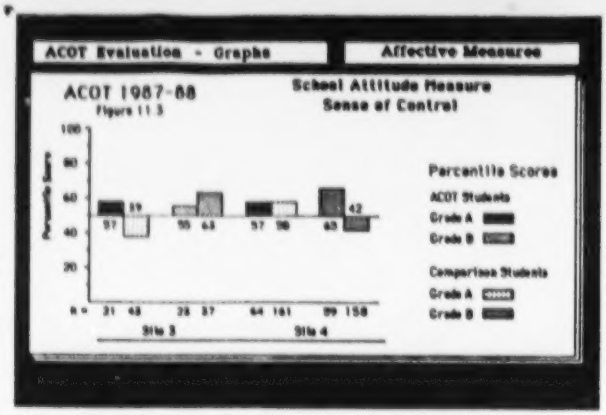

ACOT Evaluation - Graphs
Affective Measures
ACOT 1987-88
Figure 10.3
School Attitude Measure
Sense of Control

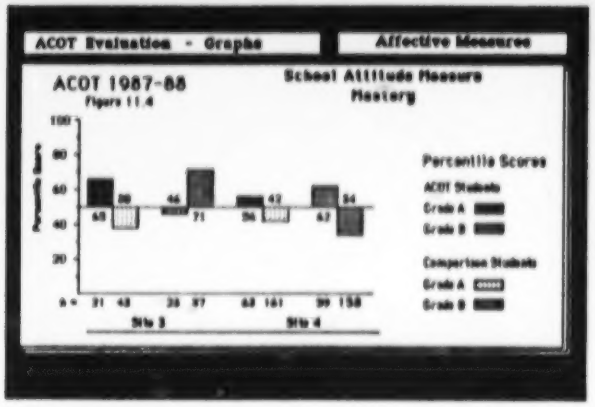

ACOT Evaluation - Graphs
Affective Measures
ACOT 1987-88
Figure 10.4
School Attitude Measure
Mastery
Percentile Scores

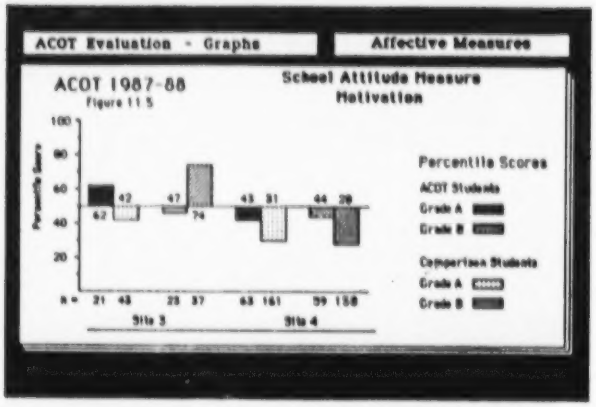

ACOT Evaluation - Graphs
Affective Measures
ACOT 1987-88
Figure 10.5
School Attitude Measure
Motivation
Percentile Scores

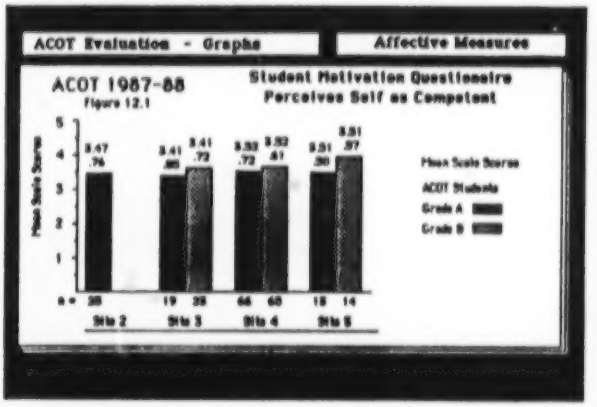

ACOT Evaluation - Graphs
Affective Measures
ACOT 1987-88
Figure 11.1
School Attitude Measure
Self Concept: Performance Based
Percentile Scores
Site 3
Site 4

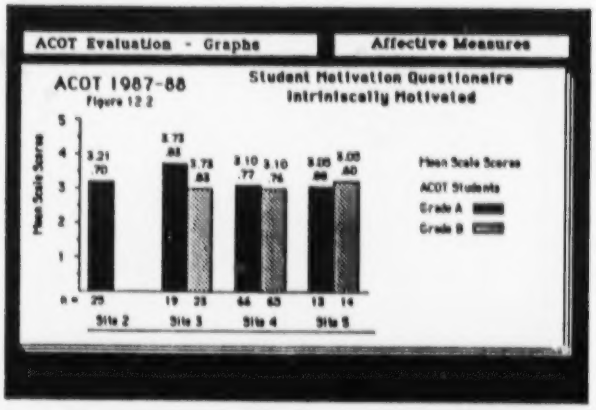

ACOT Evaluation - Graphs
Affective Measures
ACOT 1987-88
Figure 11.2
School Attitude Measure
Self Concept: Reference Based
Percentile Scores
Site 3
Site 4

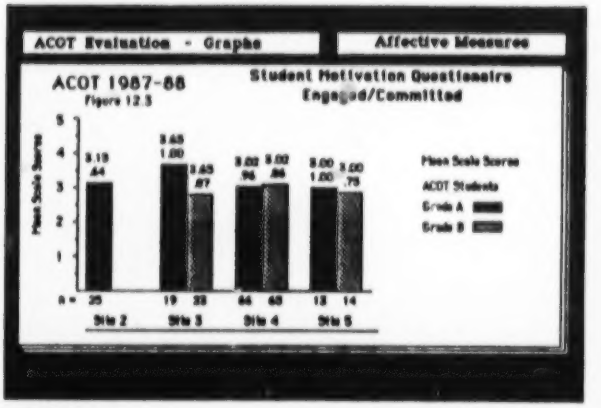




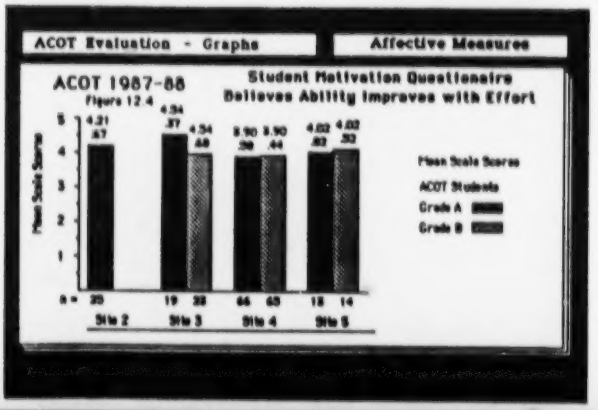


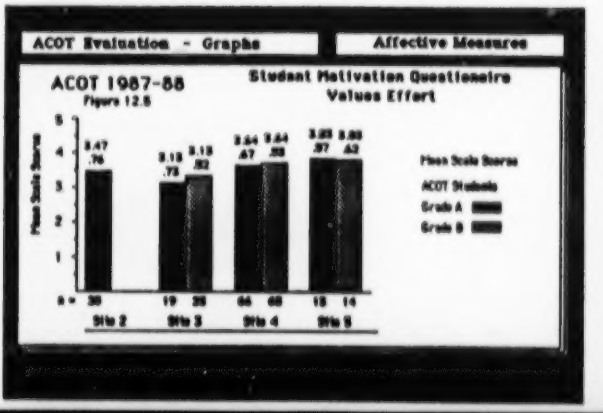


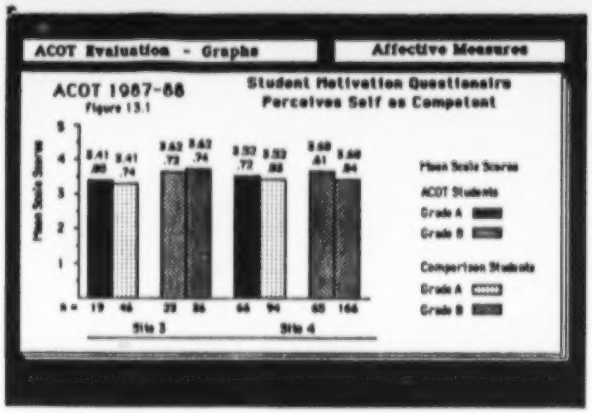


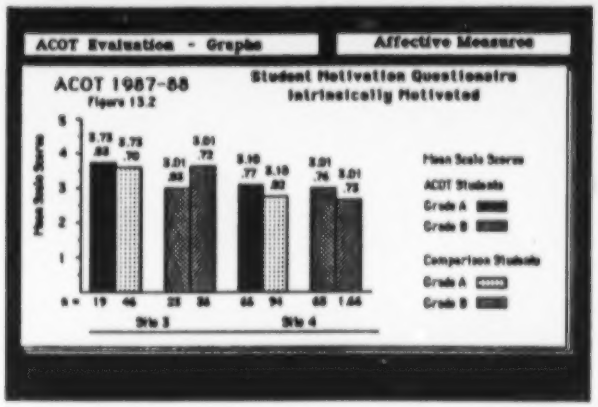


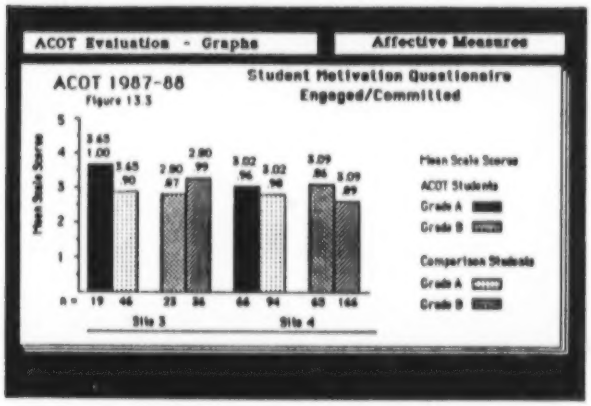
BEST COPY AVAILABLE

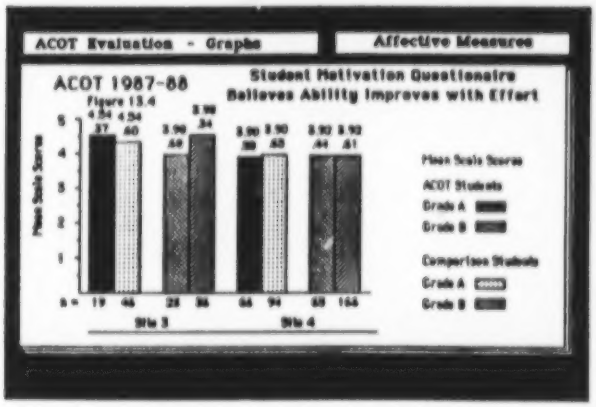


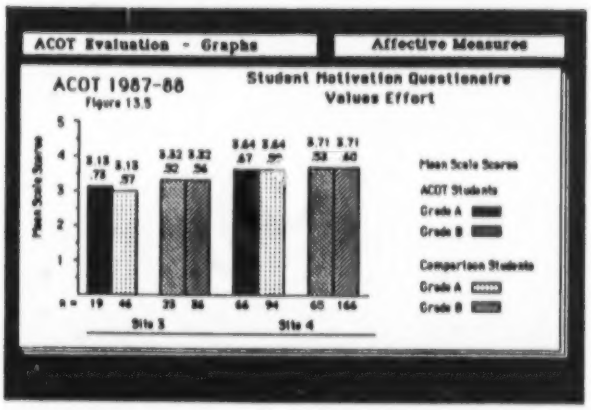












BEST COPY AVAILABLE